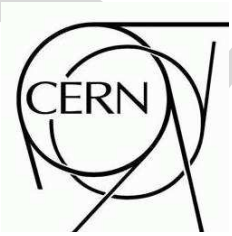
Draft version 1.0

ATLAS NOTE

ATL-PHYS-PUB-2006-000

October 26, 2007



A pre-commissioning $t\bar{t}$ cross section measurement at ATLAS

A. Bach, A. Gaponenko, I. Hinchliffe, A. Holloway-Arce,
P. Loscutoff, M. Shapiro, L. Skinnari, L. Tompkins, J. Virzi

Abstract

We describe the measurement of the $t\bar{t}$ cross section in the electron + jets channel, using the inclusive electron streaming dataset. Using events passing the LVL2 25 GeV isolated electron trigger, we observe 486 $t\bar{t}$ candidate events with a tight electron and missing E_T (consistent with a $W \to e\nu$ decay) and four or more central jets in the data. After correcting for electroweak, diboson, and single top background sources, we find that this corresponds to a $t\bar{t}$ cross section of FINAL COUNTING RESULT. Fitting the distribution of the number of jets in the sample of W candidates gives an estimate of the $t\bar{t}$ cross section of FINAL FITTED RESULT. We also describe studies of the number of b-tagged events and of events with a tight muon from the same trigger, which are consistent with the measured top cross section and could be used to refine the analysis.

Contents

1 Introduction **3**

2 Data selection **3**
2.1 Object definitions . . . 3
2.2 Event selection . . . 5

3 Calibrations and efficiencies **7**
3.1 Electron energy scale calibration . . . 7
3.2 Missing transverse energy scale . . . 8
3.3 Trigger Efficiency . . . 8

4 Signal acceptance **9**
4.1 Acceptance calculation . . . 9
4.2 Signal modeling systematics . . . 9
4.3 Effect of energy scale uncertainties on the signal acceptance . . . 10

5 Backgrounds **11**
5.1 Backgrounds with a real electron . . . 11
5.1.1 Single Bosons . . . 12
5.1.2 Single Top . . . 12
5.1.3 Gauge Boson Pairs . . . 13
5.2 Detector backgrounds . . . 13
5.3 Systematic uncertainties on the background estimates . . . 13
5.3.1 W jet multiplicity using Z . . . 13

6 Cross section results **13**
6.1 Counting method . . . 13
6.2 Fitting method . . . 15

7 Results **15**

8 Refinements and other analyses **15**
8.1 *b*-Tagging . . . 19
8.2 Dilepton Mode . . . 19
8.3 Reconstructed Top Mass . . . 22

9 Conclusion **23**

A Transfer function **23**
A.1 Truth jet transfer function . . . 23
A.2 Electron transfer function . . . 25
A.3 Transfer function systematics . . . 25

B MC samples **25**
B.1 Samples for $t\bar{t}$ signal . . . 25
B.2 Samples for electroweak backgrounds . . . 27
B.3 Samples for single top and dibosons . . . 27

1 Introduction

Motivation+Method

Define data=streaming data, MC = MC.

2 Data selection

We use the inclusive electron streaming dataset, generated from a mixture of physics processes simulated in release 11.0.42 and reconstructed with release 12.0.6.5. The dataset corresponds to a nominal luminosity of 18 pb^{-1}. The streaming event generation includes simulated online "dead-time" and some luminosity blocks of bad data. Using the prototype luminosity/conditions database [**?**] to account for deadtime corrections and file losses, the luminosity in the inclusive electron samples[1)] is 15.03 pb^{-1}. Removing the four luminosity blocks marked "BAD" in the database, we are left with 14.86 pb^{-1} of data for this analysis.

2.1 Object definitions

This section describes our object-level cuts that define what we call an "electron," a "jet" and missing transverse energy; then describes the event-level cuts that we use in this study.

Electron definition

An electron is an object from a `ElectronContainer` with the StoreGate key *ElectronCollection*, which satisfies:

1. AuthorEgamma,
2. $|\eta| < 2.4$ and $|\eta| \notin [1.37, 1.52]$,
3. $p_T > 25$ GeV.
4. `isEM&0x7FF == 0`

Distributions of the electron p_T and η are shown in Fig. 1.

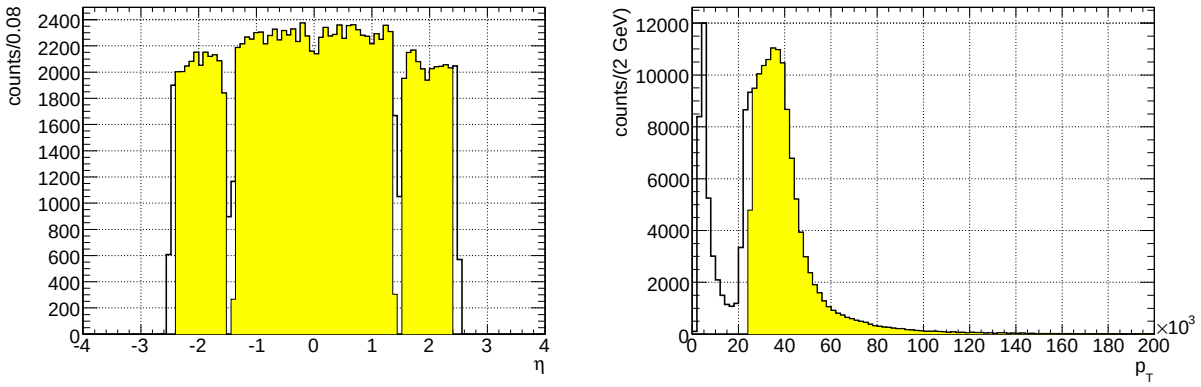
Jet definition

A jet is an object from an `ParticleJetContainer` with the StoreGate key *Cone4TowerParticleJets*, which satisfies:

1. $|\eta| < 2.5$
2. $p_T > 25$ GeV
3. To avoid double counting an electron as a jet, jets that are close to electrons passing the cuts described above are removed by the requirement $dR(\text{electron}, \text{jet}) > 0.3$

Note that no default overlap removal is performed between jets and μ, τ or photon candidates, since these objects are not used in this analysis. For the discussion of the dilepton mode in Section 8.2, where a muon candidate is required, the muon is required to be separated from a jet by $dR(\mu, jet) > 0.3$, so no overlap removal is necessary. Distributions of variables used in overlap removal are shown on Fig. 2, p_T and η cut variables on Fig. 3.

[1)] These are the ten datasets `streamtest.00*.inclEle.merge.AOD.v12000605`.



(a) (b)

Figure 1: Distributions of η (a) and p_T (b) of AuthorEgamma electron candidates. The open histogram for the η distribution includes only those candidates that passed the p_T cut, and for p_T distribution only the candidates that passed the η cut. Solid histograms are for electrons after both η and p_T cuts.

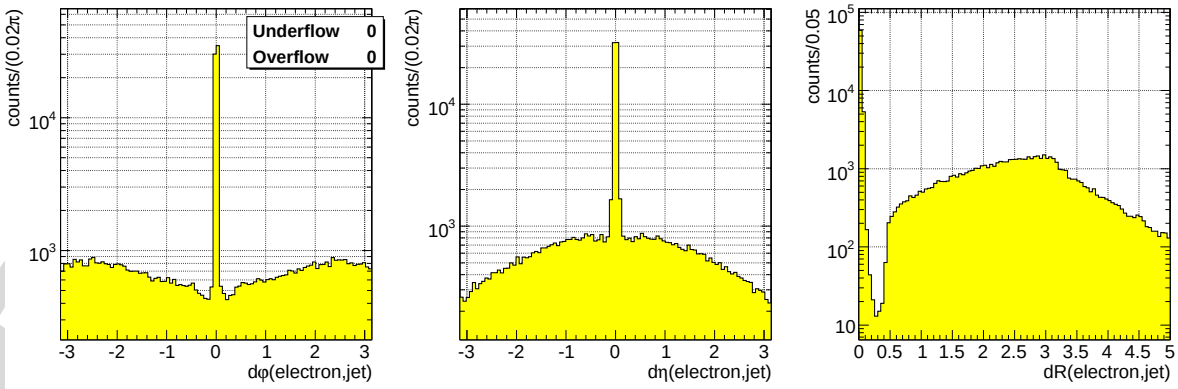
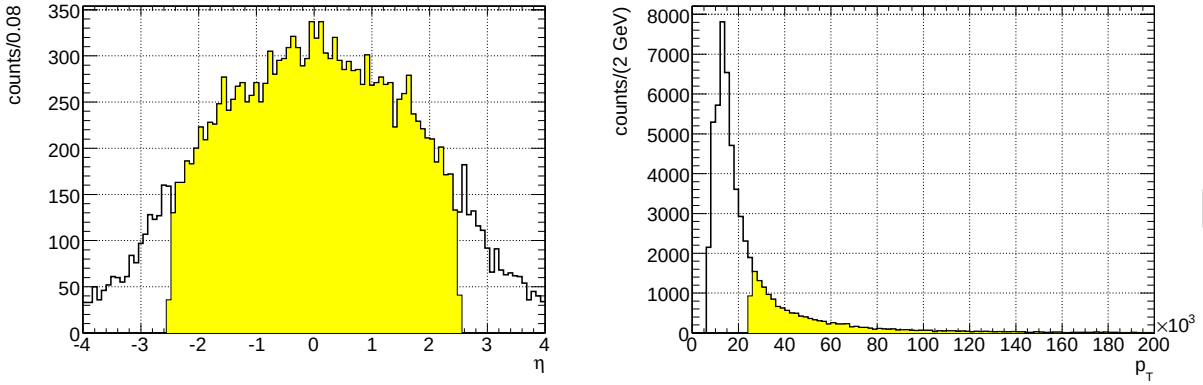


Figure 2: Distributions of jet-to-electron distances in η, φ, and R, before jet-electron overlap removal is performed. Entries in these plots use electrons after all cuts are applied, and jet candidates before any cuts are applied.



(a) (b)

Figure 3: Distributions of η (a) and p_T (b) of jet candidates, after electron-jet overlap removal. The open histogram for the η distribution includes only those candidates that passed the p_T cut, and for the p_T distribution only the candidates that passed the η cut. The solid histograms are for jets after both η and p_T cuts.

Missing Transverse Energy

The $\not\!\!E_T$ is obtained from the `MissingET` object with StoreGate key *MET_RefFinal*.

2.2 Event selection

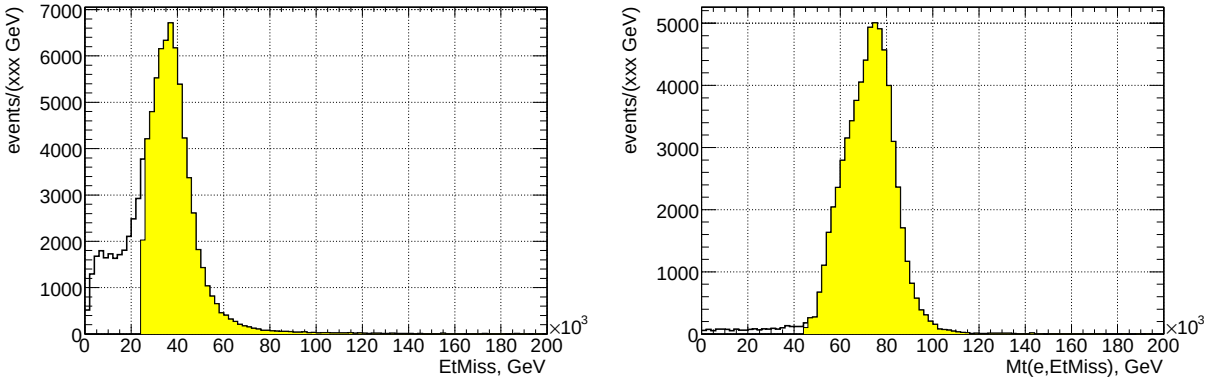
Event selection cuts are designed to obtain an inclusive sample of $W \to e\nu$ events.

- Events are required to pass the L2_e25i trigger.
- An event must have exactly one electron, as defined above. The electron requirements were imposed in stages, as shown in Fig. 5. An event is said to fail the "electron author" cut if there are no AuthorEgamma electron candidates in the input collection. Similarly, if there are no input electrons inside the accepted η or p_T range, or none passes the isEM requirement, the event fails corresponding cut in Fig. 5.
- $\not\!\!E_T > 25$ GeV, the distribution of this variable before and after the cut is shown on Fig. 4(a)
- Transverse W mass cut

$$m_t = \left(E_T(e) + \not\!\!E_T\right)^2 - \left(\vec{p}_T(e) + \vec{\not\!\!E}_T\right)^2 > 45 \text{ GeV} \quad (1)$$

  The distributions before and after the cut are shown in Fig. 4(b)
- In addition to the trigger bit requirement, we require that the reconstructed electron matches a trigger electron that passes the L2_e25i cuts. A match means $dR < 0.2$. This matching requirement is necessary in order to be able to measure trigger efficiency using a tag and probe method.

The cuts above define the inclusive W selection. For the counting analysis discussed in section **??**, the $t\bar{t}$ sample is defined by the final cut shown on Fig. 5, which requires a minimum of 4 jets.



(a) (b)

Figure 4: Distributions of cut variables $\not\!\!E_T$ and $m_T(e, \not\!\!E_T)$.

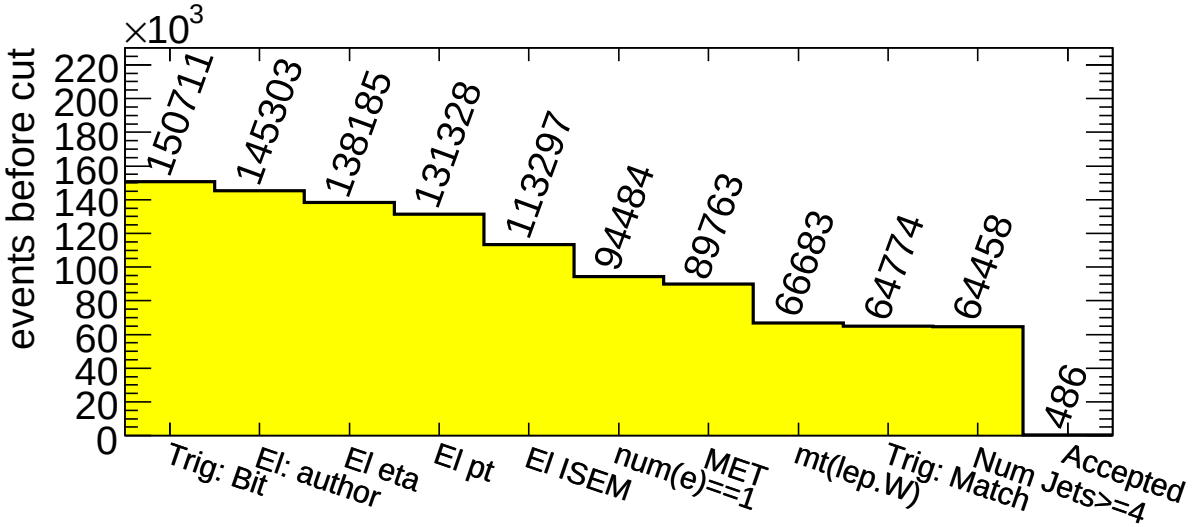


Figure 5: Number of events before each cut, in the streaming dataset. The inclusive $W \rightarrow e\nu$ sample contains 64458 events. The very last cut (Num Jets ≥ 4) is only applied in the counting analysis.

3 Calibrations and efficiencies

The stream test data used in this analysis was simulated using release 11 but reconstructed using release 12. As a result, the calibration factors applied during reconstruction are not optimal. In addition, the Monte Carlo samples used to measure the acceptance have been simulated with release 12, which has more material than release 11. We have chosen to handle these differences using a technique similar to that which would be applied to real data. We treat the streaming data as our experiment and the release 12 simulation as imperfect Monte Carlo. By comparing the two, it is possible to derive corrections. We correct the Monte Carlo as needed. We also measure the trigger efficiency using the streaming data itself.

3.1 Electron energy scale calibration

For electrons, we correct the electromagnetic energy scale of the release 12 Monte Carlo to agree with the scale observed in the streaming data. Before correction, a miscalibration is evident in the different shapes of the Z mass peak in streaming data and in a PYTHIA $Z \to ee$ sample[2)], as shown in Figure 6.

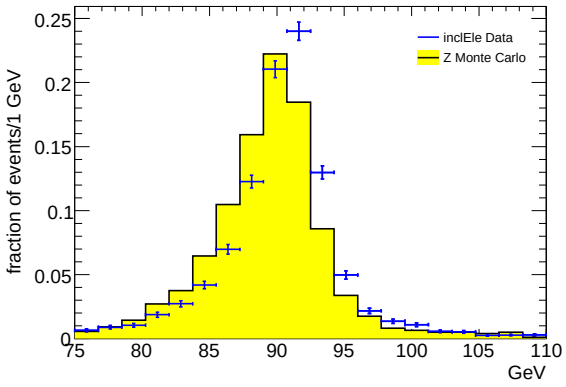


Figure 6: Differences in the electron energy scale in streaming data and the release 12 simulation sample show a systematically shifted reconstructed Z mass.

We assume that the effect of miscalibration can be represented by a factor that is a product of independent functions of electron pseudorapidity and energy, so the corrected energy can be written $E_{\rm corr} = \alpha_1(\eta)\alpha_2(E) \cdot E_{\rm sim} \equiv \alpha(\eta, E) \cdot E_{\rm sim}$. We then determine the correction factor $\alpha(\eta, E)$ by calibrating the mean Z mass in bins of η or E. The value of M_Z^2 reconstructed using corrected electron and positron energies is $\alpha(\eta_+, E_+)\alpha(\eta_-, E_-)M_{\rm sim}^2$. To measure the correction, we equate this to the mean value of M_Z^2 from the streaming data.

In Figure 7, the average value of M_Z^2, scaled to the world average, is represented as a function of the lepton's energy and pseudorapidity for the streaming data and the release 12 simulation. The data distributions, proportional to $\alpha(\eta_\pm, E_\pm)\langle\alpha(\eta_\mp, E_\mp)\rangle$ for positrons (electrons), have no discernible dependence on the charge of the lepton. We combine the electron and positron plots to derive the calibration: the result is shown in Figure 8. The η and E distributions are consistent with a constant correction factor of 1.009 ± 0.001 in the range ($E > 25$ GeV) and ($0 < |\eta| < 1.3$ or $1.7 < |\eta| < 2.4$). We treat the variation of the correction in the cracks near $|\eta| = 1.5$ as a systematic uncertainty.

We may incur a systematic bias by assuming that the correction is independent of electron energy. Allowing a linear term in the fit to $\alpha(E)$, the correction varies by $^{+0.002}_{-0.001}$ in the range 25 to 101 GeV

[2)] We use `trig1_misal1_mc12.005144.PythiaZee.recon.AOD.v12000604`.

(which encompasses 90% of the leading electrons in selected events in the $t\bar{t}$ simulation). If we include the regions near the crack, ($1.3 < |\eta| < 1.7$), the derived correction shifts by 0.004. We therefore combine a systematic error of 0.002 with the statistical error on the fit, so that the electromagnetic energy scale is known with a 0.22% relative uncertainty.

3.2 Missing transverse energy scale

The missing transverse energy used to select W candidates in this analysis is calculated from a sum over specifically calibrated calorimeter cells in three categories: cells in electromatic clusters, in jets, and in clusters not associated with any reconstructed calorimeter object [1]. This sum is then corrected for the E_T of identified muon candidates and for probable energy loss in the cryostat. Since the cell energies recieve either electromagnetic or hadronic energy scale corrections, a systematic miscalibration of the $\not\!\!E_T$ could result from miscalibrations of either scale, or of the muon identification efficiency.

As a first comparison of the scale of missing energy measurements in the release 12 simulation and the streaming data, we analyze the W transverse mass distribution. This distribution is unaffected by differences in the W boson kinematics, but other sources of true missing energy such as additional neutrinos or unidentified muons will distort this distribution in the streaming data. We use the inclusive W cuts described in section 2.1 to select W candidate events in the streaming data and a PYTHIA $W \to e\nu$ sample[3)] simulated in release 12.0.6. We apply the lepton energy scale correction derived in section 3.1 and subtract the change in the electron's transverse momentum vector from the missing energy. By requiring that the multiplicity of jets with p_T greater than 25 GeV be less than two we exclude most $t\bar{t}$ events. The W transverse mass reconstructed in each sample is plotted in Figure 10. The ratio between the mean $M_T(W)$ in the streaming data and the PYTHIA sample is 0.985 ± 0.001 when $N_{\text{jets}} = 0$ and 0.973 $\pm$ 0.003 when $N_{\text{jets}} = 1$. This difference suggests a possible dependence of the $\not\!\!E_T$ scale on the amount of jet activity. Because the $t\bar{t}$ events have a higher level of jet activity then the W sample, extrapolation of the $\not\!\!E_T$ scale is difficult. We therefore choose to retain our uncorrected $\not\!\!E_T$ scale and to assess a systematic uncertainty due to any miscalibration.

3.3 Trigger Efficiency

We measure the electron trigger efficiency (with respect to reconstruction) of the L2_e25i trigger by applying the tag and probe method to electrons in the $Z \to ee$ peak. "Reconstructed" electrons are those that passed the cuts mentioned in Section 2.1, including all of the isEM cuts except TRT. In events where there are two good electrons (of opposite charge) that give an invariant mass of $m_Z \pm 10$ GeV, we apply the standard tag and probe procedure [3] and plot the trigger efficiency with respect to p_T and η in Fig. 11. The trigger efficiency ε_t is given by:

$$\varepsilon_t = \frac{2N_2}{N_1 + N_2}$$

and the statistical uncertainty is given by

$$\sigma_{\varepsilon_t} = \sqrt{\frac{\varepsilon_t(1-\varepsilon_t)}{N_1 + N_2}}$$

where N_1 is the number of events with at least one electron passing the trigger and N_2 is the number of events with two electrons passing the trigger. Since the distribution is essentially flat for $p_T > 25$ GeV and for η outside the cracks, we quote a simple overall trigger efficiency rather than convoluting it with the p_T distribution of the electrons.

[3)] We use `trig1_misal1_csc11.005100.JimmyWenu.recon.AOD.v12000601`, applying the "1mm" bug correction in the AOD [2].

For electrons that pass our p_T and η cuts, we find the trigger efficiency with respect to reconstruction to be $\varepsilon_t = (98.96 \pm 0.11)\%$. The error quoted is statistical. The primary background to events passing our Z mass window cut are W and top events with two good electrons. The properties of such electrons, as far as the trigger is concerned, should be identical to those of $Z \to ee$ electrons. Thus, such background does not introduce a systematic bias in the efficiency.

One source of systematic uncertainty, however, arises from a peculiarity of the streaming data. Between the time when the trigger code was run to to create the streamed data (and hence used to create the trigger mask in the event header) and the time when the data were reconstructed (including reconstruction of trigger objects), the trigger code changed, affecting the overall trigger efficiency. In the one typical run of the electron streaming data, we observe 14626 events where the trigger was satisfied based on the TriggerDecision created during reconstruction and 14689 events where the relevant bit was set in the event header. A total of 14471 of these events were in common. Since our tag-and-probe method requires that the trigger be satisfied in the reconstruction but our original events were selected using the streaming bit, we assign an uncertainty on the electron trigger efficiency.

4 Signal acceptance

4.1 Acceptance calculation

In this section we present the acceptance of our event selection for $t\bar{t}$ signal events generated with MC@NLO. The MC@NLO sample used was created with a generator-level filter requiring a lepton with $p_T > 5$ MeV, so in this note we refer to efficiencies with respect to this $t\bar{t}$ inclusive lepton sample. The efficiency of our event preselection for $t\bar{t}$ leptonic events is $0.13 \pm$ XXX and the final efficiency, which includes the jet multiplicity cut, is $0.053 \pm$ XXX, where the errors quoted are statistical. We explore systematic errors on the signal acceptance in Section 4.2.

Event selection requirement	**acceptance (relative to previous cut)**
Generator filter (single lepton, $p_T > 5$ MeV)	0.554 ± 0
AuthorEgamma electron	0.9572 ± 0.0013
electron η requirements	0.9874 ± 0.0007
electron $p_T \geq 25$ GeV	0.6002 ± 0.0032
electron isEM requirement	0.3565 ± 0.0041
exactly one electron	0.9695 ± 0.0025
$\not\!\!E_T \geq 25$ GeV	0.8605 ± 0.0051
W $m_T > 45$ GeV	0.7712 ± 0.0065
$N_{\text{jets}} \geq 4$	0.4070 ± 0.0087

Table 1: The acceptance of our event selection (excluding trigger requirements) for signal events. MC@NLO weights are used for all computations.

4.2 Signal modeling systematics

Monte Carlo generator

We use MC@NLO [4] version 3.1, with Jimmy [5] showering, to generate the $t\bar{t}$ signal events and determine our acceptance. This generator includes all terms in the matrix element up to order α_s^3, but neglects some observable angular correlations. As a very crude estimate of the theoretical uncertainty,

Generator	**acceptance of inclusive W cuts**	**acceptance of $t\bar{t}$ cuts**
ACERMC with duplicate events	$0.127 \pm$?	$0.058 \pm$?
MC@NLO	$0.130 \pm$	$0.053 \pm$
PYTHIA	0.106 ± 0.4	0.053 ± 0.003

Table 2: The acceptance of our inclusive W and $t\bar{t}$ event selection (excluding trigger requirements) for events generated with ACERMC, MC@NLO, and PYTHIA. **Because of a duplicated event problem in the** ACERMC **sample, the statistical errors cannot be reliably calculated. This table will be updated when a corrected sample is available.**

Sample	PYTHIA **settings**	**acceptance**
AcerMC with "low m_T" PYTHIA settings	PARJ(81) = 2× default PARP(61) = default ÷ 2 PARP(62) = 2× default	0.063 ± 0.001
AcerMC with default PYTHIA settings	PARJ(81)= 0.25 GeV PARP(61)=0.192 GeV PARP(62)=1 GeV	$0.058 \pm$?
AcerMC with "high m_T" PYTHIA settings	PARJ(81) = default ÷ 2 PARP(61) = 2× default PARP(62) = default ÷ 2	0.052 ± 0.001

Table 3: Variation in the signal acceptance from the ACERMC + PYTHIA generators when parameters are changed to explore the uncertainty due to initial and final state radiation.

we compare the acceptance calculated in section 4.1 to the acceptance derived with PYTHIA alone, and to the acceptance derived with ACERMC[4)].

Initial and final state radiation

Uncertainty in the modeling of initial and final state radiation affects the average number of jets above threshold in top events, and thus the acceptance of our event selection (especially the final $N_{\text{jet}} \geq 4$ requirement). In Table 3, we compare the $t\bar{t}$ event acceptance calculated with three different PYTHIA configurations.

4.3 Effect of energy scale uncertainties on the signal acceptance

Electromagnetic energy scale calibrations

The calibration in Section 3.1 resulted in a systematic uncertainty of about 0.22% on the corrected electron energy scale. The acceptance for $t\bar{t}$ signal events when the electron energy correction is varied by

[4)] Due to a production job configuration error, the *same* input events were simulated and reconstructed many times in our ACERMC $t\bar{t}$ dataset. Hence, the statistical error on the acceptance for this sample is not known, and the results will be updated when a new ACERMC $t\bar{t}$ sample has been processed.

$\pm$ 0.2% (and the $\not\!\!E_T$ and W transverse mass are recalculated) scale does not change appreciably.

Jet energy scale

We cannot calibrate the jet energy scale using the inclusive electron streaming data. With real data, this calibration could be performed using photon-jet balancing and related studies [6]. We check the sensitivity of the signal event acceptance to the nominal 5% uncertainty which should eventually be achievable for generic jets. The change in acceptance is about 6%.

jet energy scale	acceptance of $t\bar{t}$ selection	relative change
1.05	5.55 ± 0.15 %	+6%
0.95	4.97 ± 0.14 %	-5%

Table 4: The acceptance of our event selection (excluding trigger requirements) with varied jet energy scales.

$\not\!\!E_T$ scale uncertainty

The study in Section 3.2 indicates a systematic uncertainty of about 3% for low jet multiplicities. However, the method is too sensitive to contamination by $t\bar{t}$ events to be used in the high jet multiplicity region, and comparision of the ratios determined in the 0- and 1- jet bins does not rule out a correlation with jet activity. To estimate a systematic uncertainty, we simply assume that the jet energy scale miscalibration is the dominant driver of the missing energy scale in events with a large jet multiplicity. We therefore assign the $\not\!\!E_T$ scale the same nominal 5% uncertainty as the jet energy scale and calculate the effect of such an uncertainty on the signal acceptance, which is shown in Table 5. The resulting systematic uncertainty is 4%.

Missing E_T scale	Lepton+jets acceptance	relative change
1.05	0.054 ± 0.001	+4%
0.95	0.050 ± 0.001	-4%

Table 5: The acceptance of our event selection (excluding trigger requirements) with different missing energy scale settings.

5 Backgrounds

In additon to top pair production and decay, several other processes contribute to the inclusive W sample and the final $t\bar{t}$ sample. These are of two types, events having a real isolated electron that pass the trigger and events with the trigger is a fake electron arising from detector effects.

5.1 Backgrounds with a real electron

The most important background comes from events with a real W boson, that decays into electron either directly or via a τ lepton. Background from Z+jets events where one of the electons is lost, is small because the missing transverse energy cut rejects it. The single top production cross section at the LHC is significant, therefore single top decays constitute a non negligible background. Backgrounds from dibosons (WW, WZ, ZZ) are smaller as the production cross sections of these processes are small.

Backgrounds are analysed using release 12 Monte Carlo samples. Since the trigger efficiency was determined from the stream test data in section 3.3 and is independent of lepton p_T, trigger information in the Monte Carlo samples was not used. the data is corrected by a constant factor. We applied this factor to reconstructed electrons in Monte Carlo samples and corrected energy of as described in section 3.1.

In the following subsections we describe the backgrounds and their systematic uncertainties. Table 6 summarizes the results.

MC sample	Monte Carlo cross section [pb]	Filter efficiency	Number of input events	Acceptance of the inclusive W selection	Predicted number of events in 14.86 pb^{-1}
5104 $W \to e\nu$	17440	0.625	435750	0.3412 ± 0.0007	55265.53 ± 113.38
5106 $W \to \tau\nu$	17170	0.198	153350	0.0218 ± 0.0004	1101.31 ± 20.21
5144 $Z \to ee$	1675	0.855	14700	0.0050 ± 0.0006	106.41 ± 12.77
5500 single top Wt	26.7	1	48350	0.1092 ± 0.0014	43.33 ± 0.56
5501 single top s-channel	3.3	1	48300	0.0884 ± 0.0013	4.33 ± 0.06
5502 single top t-channel	81.3	1	44450	0.1003 ± 0.0014	121.17 ± 1.69
5985 WW	70	0.35	50000	0.1572 ± 0.0016	57.23 ± 0.58
5987 WZ	27	0.29	47900	0.1201 ± 0.0015	13.97 ± 0.17
5986 ZZ	11	0.19	49800	0.0373 ± 0.0008	1.16 ± 0.02

Table 6: Summary of MC samples used for background estimates in central value extraction. Full names of the samples and their AMI provenances are listed in Appendix B. The errors in the last column include statistical errors on the acceptance only. The difference between $W \to e\nu$ and $W \to \tau\nu$ cross sections indicates the size of statistical error on these MC cross sections. Cross section and filter efficiency numbers are taken from ATLAS Wiki pages `WZPythiaSample`, `TopDC3`, and `GaugeBosonPairs`.

5.1.1 Single Bosons

There are considerable differances in the total rates and in the jet multiplicities in the Monte-Carlo estimates [7]. The uncertainties in the absolute rates are not relavent to this analysis as we normalized them to the data. The primary systematic uncertainty arises from the jet multiplicity. As the baseline, we use Pythia generated and simulated with release 12 (Pythia version XXXXX): samples 5104, 5106 and 5144 (see appendix for details). In addition, we use pythia (5104) generated with release 11 (pythia YYYYY) and simulated with release 12 and Herwig/Jimmy (5100) generated with release 11 and simulated with release 12.

To further study the systematics we generated additional samples varying Monte Carlo parameters as described in the appendix. Fully simulating these additional sample is prohibitive. Atlfast *add ref* does not include efficiencies for electrons or jets and is not calibrated to these data sets. Therefore it was not used. Instead, we dervied a transfer funtion and appied it to truth electrons and truth jets: truth jets and reconstructed jets use the same jet algorithm. Details of this proceedure are in appendix *fix this*

5.1.2 Single Top

There are three processes that contibute here. The s-channel process resulting in a $t\bar{b}$ final state and two t-channel processes resulting in tW or $t\bar{d}$ and $t\bar{s}$. These are simulated using AcerMC [8]. The s-channel process is unimportant as can be seen from table 6. In the fitting method discussed below the magnitude

of the t-channel processes is allowed to float and s-channel is neglected. The jet multiplicity distributions of the two t-channel processes is similar and is constrained from the Monte Carlo. Systematics have not been fully investigated. Uncertainties in the total rate using MCVNLO is approximately 0.75 of that from AcerMC [9].

5.1.3 Gauge Boson Pairs

5.2 Detector backgrounds

In real data, we anticipate that jets faking electrons will be a substantial background. However, since jet rejection is of the order 10^3, simulating a large number of these fake electrons is computationally prohibitive. Thus only a very small number exist in the streaming data. To verify that this background is negligible, we plot the distributions of the electron identification variables used in isEM. See Fig. 13 [MAKE LOG] for example plots from data. As expected, the tails of the distributions for all variables are extremely small, both for all electrons and for electrons that have passed other purification cuts.

As expected, the background of fake electrons is negligible. In real data, distributions such as these could be used to quantitatively assess the background by extrapolating the tails intothe acceptance region.

5.3 Systematic uncertainties on the background estimates

5.3.1 W jet multiplicity using Z

It is essential to have a good measurement of jet multiplicity in $W \to e\nu$ events, because this process is the largest background to $t\bar{t}$ events. This multiplicity cannot be obtained directly from data because the high multiplicity bins are contaminated by $t\bar{t}$ events. We can, however measure the jet multiplicity in $Z \to ee$ events, which has no similar contamination. By computing the ratio of jet multiplicities between $Z \to ee$ and $W \to e\nu$ via Monte Carlo, we can use a measured jet multiplicity for $Z \to ee$ events to estimate the jet multiplicity in $W \to e\nu$ events in our background. Fig. 14 shows this ratio. The two processes do not have identical distributions of jet multiplicity, but the difference is well behaved and can be explained via the Sudakov factor.

Z vs W. Can use Z in higher Nj bins to estimate the background without being affected by ttbar. But need W/Z ratio and systematic on it.

6 Cross section results

We extract $t\bar{t}$ cross section results from the jet multiplicity distribution of the inclusive $W \to e\nu$ event sample defined in section 2.2. The two approaches we use to calculate $t\bar{t}$ cross section are described below. Table 7 shows normalized jet multiplicity spectra for streaming data, an MC@NLO $t\bar{t}$ signal sample, and background samples.

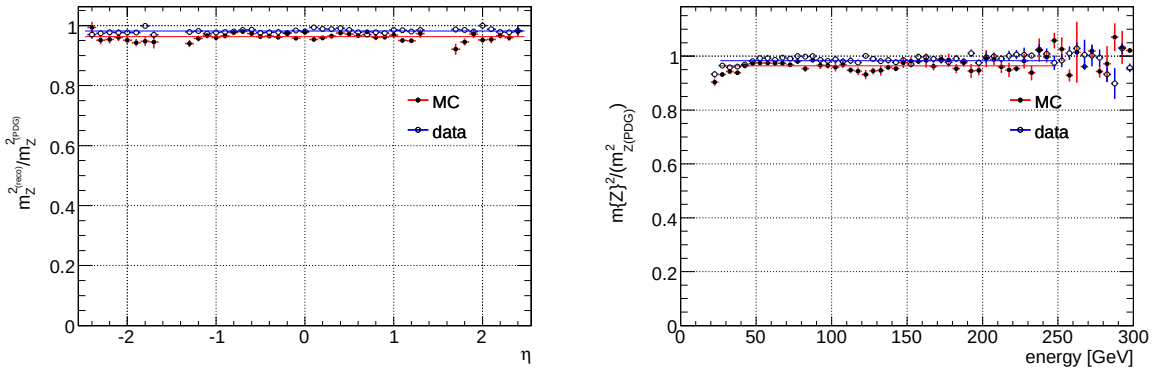
Sanity check: extracting W cross section.

Table of multiplicities for streaming, ttbar, backgrounds.

6.1 Counting method

The "counting" method is similar to the method used by CDF for the first top observation [**?**]. We use the inclusive W event selection described in section 2.2, and impose an additional cut on jet multiplicity (at least 4 jets) to select top-enriched subsample. We estimate background normalization by using events with 0 and 1 jets, which contain a negligible amount of top decays.

The dominant background is the $W \to e\nu$ process, therefore we scale cross section for this sample to match the number of events in 0+1 jet bins observed in streaming data. That means that $W \to \tau\nu$ cross

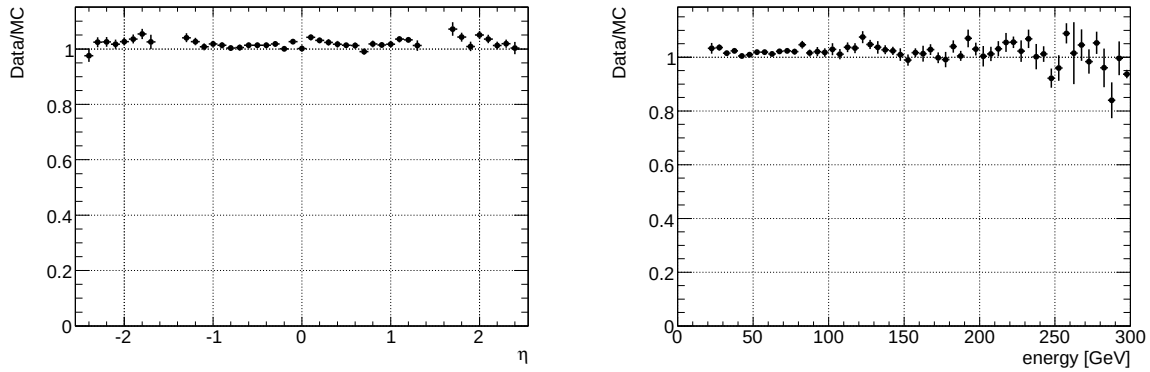


(a) Z mass dependence on electron η. (b) Z mass dependence on electron E.

Figure 7: Dependence on lepton kinematics of the reconstructed Z mass in streaming data and release 12 PYTHIA Monte Carlo. In each plot, the average Z mass squared is shown scaled by $M_Z^2(PDG)$, and each reconstructed mass makes two contributions to the profile histogram.

Sample	Fraction of accepted inclusive W events with given jet multiplicity				
	0	1	2	3	4+
streaming data	0.8254 ± 0.0015	0.1321 ± 0.0013	0.0264 ± 0.0006	0.0085 ± 0.0004	0.0075 ± 0.0003
$W \to e\nu$	0.7980 ± 0.0010	0.1613 ± 0.0010	0.0320 ± 0.0005	0.0067 ± 0.0002	0.0020 ± 0.0001
$W \to \tau\nu$	0.7722 ± 0.0073	0.1793 ± 0.0066	0.0365 ± 0.0032	0.0084 ± 0.0016	0.0036 ± 0.0010
$Z \to ee$	0.2703 ± 0.0516	0.2703 ± 0.0516	0.3243 ± 0.0544	0.1081 ± 0.0361	0.0270 ± 0.0189
single top:					
Wt	0.0102 ± 0.0014	0.1062 ± 0.0042	0.3262 ± 0.0065	0.3321 ± 0.0065	0.2253 ± 0.0057
s-channel	0.0731 ± 0.0040	0.2948 ± 0.0070	0.4288 ± 0.0076	0.1590 ± 0.0056	0.0443 ± 0.0031
t-channel	0.0498 ± 0.0033	0.2533 ± 0.0065	0.3440 ± 0.0071	0.2333 ± 0.0063	0.1196 ± 0.0049
WW	0.3145 ± 0.0052	0.3630 ± 0.0054	0.2382 ± 0.0048	0.0704 ± 0.0029	0.0139 ± 0.0013
WZ	0.3578 ± 0.0063	0.3176 ± 0.0061	0.2353 ± 0.0056	0.0730 ± 0.0034	0.0163 ± 0.0017
ZZ	0.2939 ± 0.0106	0.3827 ± 0.0113	0.2018 ± 0.0093	0.0797 ± 0.0063	0.0420 ± 0.0047
$t\bar{t}$	0.0108 ± 0.0016	0.0710 ± 0.0039	0.2173 ± 0.0063	0.2938 ± 0.0069	0.4070 ± 0.0075

Table 7:



(a) Derived energy correction as a function of electron η. (b) Derived energy correction as a function of electron E.

Figure 8: Correction to the electron energy required for the release 12 simulation.

section should be scaled by the same factor as well. The relative cross sections of W and Z production are theoretically understood much better than their absolute values [**?**], therefore we apply the same scaling to the $Z \rightarrow ee$ sample as well. Cross sections for all other background samples are fixed to their theoretical values, as listed in Table 6, and predicted numbers of events with 0+1 jet from these sources are subtracted from the number of 0+1 jet events observed in streaming data before computing the scaling factor. The results are shown in Table 9.

6.2 Fitting method

7 Results

As noted above, we observe 486 ttbar candidate events in the 14.81 pb^{-1} of good luminosity blocks in the streaming data.

** summary of systematic errors Review of above (Table)

* Cross section, given ”all-top” hypothesis Andrei

8 Refinements and other analyses

We have, in this study, focused on analyses that can be performed on the small data sample of the stream test. There are other event selections which, although less efficient, result in higher signal to background and/or in analyses with different systematics.

The requirement of tagged b-jets in the events significantly reduces the W background: indeed it was required for the first analysis done by CDF as the signal to background ratio is lower at the Tevatron. Extraction of the cross section from this requires the additional knowledge of the b-tagging efficiency which can be obtained from the data sample itself by comparing events with one and two b-tags provided that the single top component, which also contains b-jets, can be removed or measured separately. The requirement of two leptons will provide a reduction in background. We can use this channel by starting with our original event selection based on the electron and requiring a muon in the event.

The final state can be constrained in the lepton+jet analysis by studying the invariant mass of two and three jet subsystems as is done in the commissioning analysis which uses 100 pb^{-1} **??**. A sufficiently large sample can be used to provide an in situ calibration of the light quark and b-quark jet energy scales. The following provides some preliminary results using these methods.

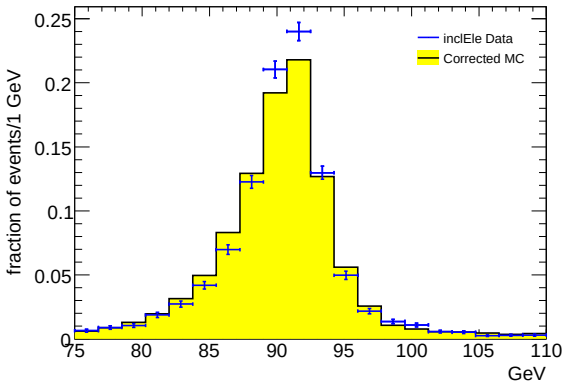
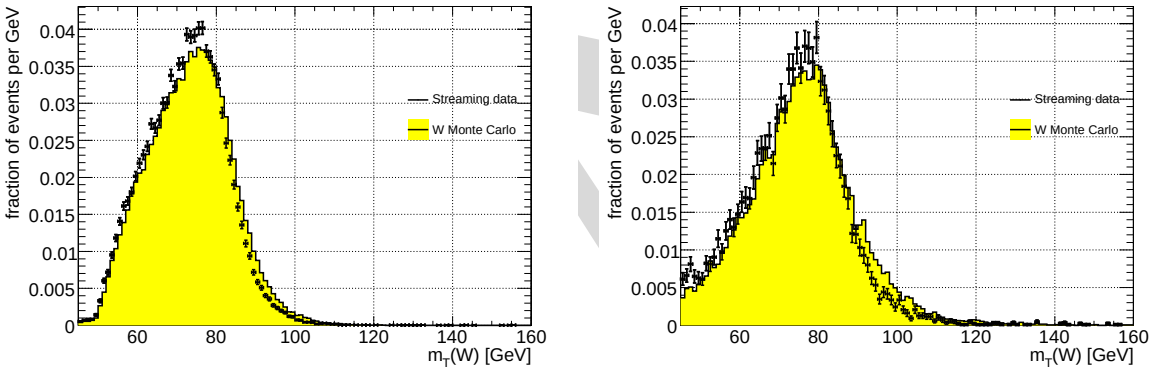
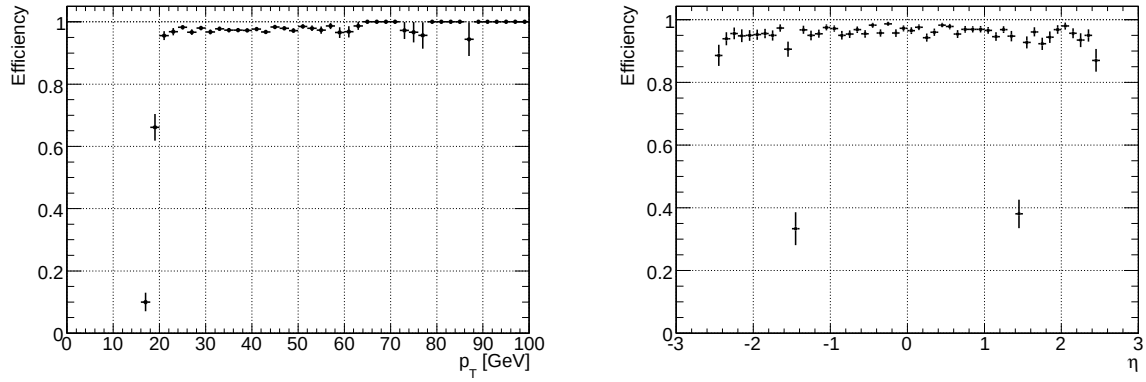


Figure 9: The reconstructed Z mass in the release 12 Monte Carlo, after the electromagnetic energy scale correction is applied, compared to the streaming data.



(a) W transverse mass for events with $N_{\text{jets}} = 0$. (b) W transverse mass for events with $N_{\text{jets}} = 1$.

Figure 10: Studies of $M_T(W)$ and the $\not\!\!E_T$ scale in streaming data and the release 12 simulation.



(a) (b)

Figure 11: The trigger efficiency for electrons from Zs, as a function of electron p_T (a) and of electron η (b). These plots are made before the p_T or η cuts are applied.

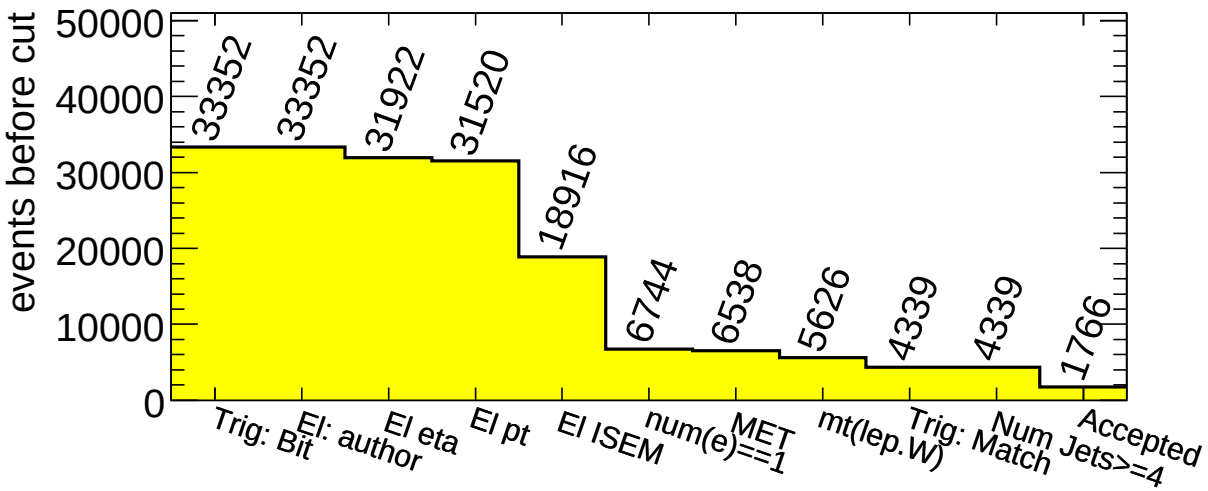
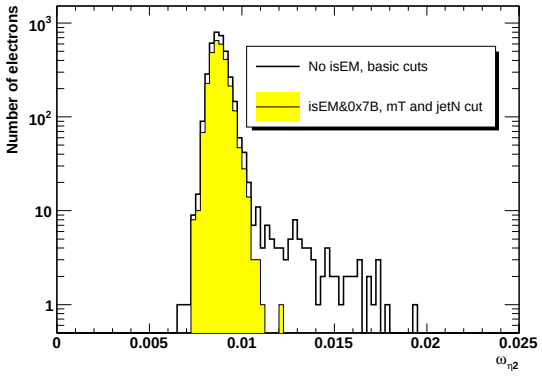
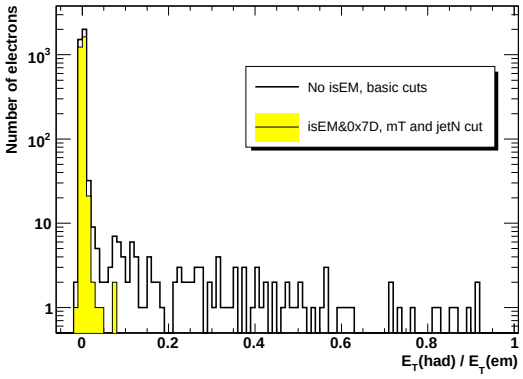


Figure 12: Graphical representation of the effect of sequential cuts on the number of events in the 005200.T1_McAtNlo_Jimmy dataset. This histogram is filled using MC@NLO weights.

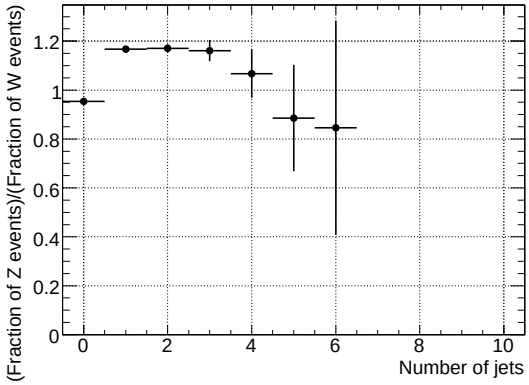


(a) The variable $\omega_{\eta 2}$, a measure of shower width in η in the second sampling of the electromagnetic calorimeter. Cut value = 0.0115.



(b) Hadronic leakage, the ratio of energy in the hadronic calorimeter to energy in the electromagnetic calorimeter. Cut value = 0.04.

Figure 13: Representative examples of electron ID variables for electrons with $|\eta| < 0.8$. They show the expected lack of fake electrons in the streaming data. Here the open histograms include all electrons before isEM cuts. The solid histograms are electrons that have passed all isEM cuts *except* the bit associated with the one being plotted, and additional event cuts to ensure a pure sample.



(a)

Figure 14: Ratio of Jet Multiplicity for $Z \rightarrow ee$ vs $W \rightarrow e\nu$ for different generators. The ratio of jets is compared after event selection cuts.

Process	Number of 0+1 jet events	Number of 4+ jet events
Streaming data	61721.00	486.00
Fixed backgrounds		
5500 Wt	5.05	9.77
5501 s-channel	1.60	0.19
5502 t-channel	36.72	14.49
5985 WW	38.77	0.79
5987 WZ	9.44	0.23
5986 ZZ	0.78	0.05
Total fixed backgrounds	92.36	25.51
Data excess over fixed bg	61628.64	460.49
Single bosons unscaled		
OLD 5104 $W \rightarrow e\nu$	54269.25	36.82
5106 $W \rightarrow \tau\nu$	1046.95	3.95
5144 $Z \rightarrow ee$	57.91	2.90
Total single bosons unscaled	55374.11	43.67
Scaling factor	1.11	
Single bosons scaled		48.60
Data excess over all bg		411.88

Table 8: OLD Calculation of the $t\bar{t}$ number of $t\bar{t}$ events in streaming data in the counting method.

Process	Number of 0+1 jet events	Number of 4+ jet events
Streaming data	61721.00	486.00
Fixed backgrounds		
5500 Wt	5.05	9.77
5501 s-channel	1.60	0.19
5502 t-channel	36.72	14.49
5985 WW	38.77	0.79
5987 WZ	9.44	0.23
5986 ZZ	0.78	0.05
Total fixed backgrounds	92.36	25.51
Data excess over fixed bg	61628.64	460.49
Single bosons unscaled		
5104 $W \to e\nu$	53015.21	110.40
5106 $W \to \tau\nu$	1046.95	3.95
5144 $Z \to ee$	57.91	2.90
Total single bosons unscaled	54120.07	117.25
Scaling factor	1.14	
Single bosons scaled		133.51
Data excess over all bg		326.97

Table 9: NEW Calculation of the $t\bar{t}$ number of $t\bar{t}$ events in streaming data in the counting method.

8.1 *b*-Tagging

Identifying jets from *b*-quark fragmentation is not necessary for isolating $t\bar{t}$ event candidates, however it is useful for calibrations and cross-checks. In this analysis we use the current default tagger 1P3D+SV1 (reference?) which is a combination of a 3D impact parameter tagger and a secondary vertex tagger. A jet is defined to be tagged as a *b*-jet if its weight is greater than 3.0.

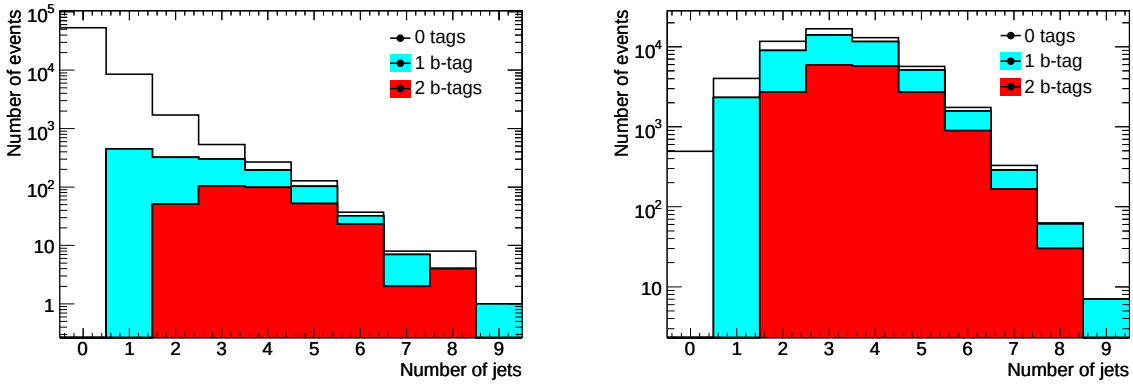
Jet multiplicity distributions for events passing the preselection cuts are plotted in Fig. 15 for streaming data and for the MC@NLO $t\bar{t}$ sample without *b*-tagging and when requiring at least one respectively at least two *b*-jets. Fig. 16 shows the sample composition of events with at least one or at least two *b*-jets. The PYTHIA electroweak background samples $W \to e\nu$, $W \to \tau\nu$ and $Z \to ee$ are, after normalization to their relative cross sections, scaled to match the number of events in the (0+1) jet bins for the streaming data without requiring *b*-tagging and this scale factor is then applied to the samples when using *b*-tagging. The $t\bar{t}$ sample and the ACERMC single top samples are all normalized according to their cross sections.

8.2 Dilepton Mode

1/9 of $t\bar{t}$ decays are fully leptonic with both W decaying into a lepton and a neutrino. The dilepton mode provides a clean sample and is despite its limited use in reconstructing the top mass valuable for cross-checks with results from the semi-leptonic mode and for providing a $t\bar{t}$ subsample with highly reduced background.

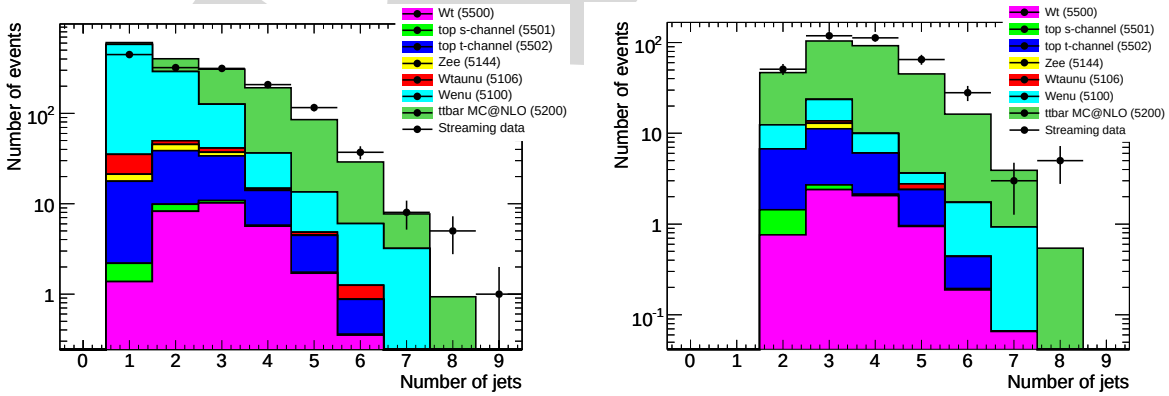
For the dilepton mode we select events with exactly one electron (as defined in section 2.1) and exactly one muon. A muon is defined as an object from a `MuonContainer` with the StoreGate key *MuidMuonCollection* and for which:

1. $p_T > 15$ GeV



(a) (b)

Figure 15: Jet multiplicity distributions for streaming data (a) and the MC@NLO $t\bar{t}$ sample (b) for semi-leptonically decaying events with zero, at least one or at least two *b*-tagged jets.

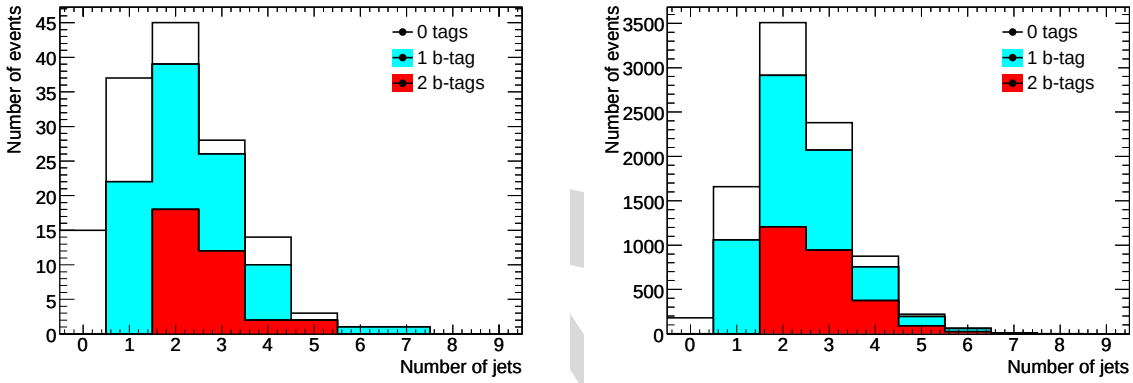


(a) (b)

Figure 16: Normalized jet multiplicities for streaming data and its different components when requiring at least one (a) or at least two *b*-jets (b) for the semi-leptonic mode.

2. $|\eta| < 2.4$
3. isolation $E_T < 6$ GeV in a cone of 0.2
4. $dR(\mu, jet) > 0.3$

For these events, jet multiplicity distributions with and without requiring b-tagging are shown in Fig. 17 for streaming data and MC@NLO $t\bar{t}$. Similarly to the semi-leptonic case, jet multiplicities are plotted for fully leptonic event candidates to show the sample composition without b-tagging (Fig. 18) and with at least one or two b-jets (Fig. 19). For the dilepton mode we include a PYTHIA $W \to \mu\nu$ sample in the electroweak background in addition to the ones used for the semi-leptonic mode. Fig. 19 only includes the background samples which had any contributions to the jet multiplicites after normalization.



(a) (b)

Figure 17: Jet multiplicities for streaming data (a) and MC@NLO $t\bar{t}$ (b) for events in the dilepton mode without b-tagging and with at least one or at least two b-tagged jets.

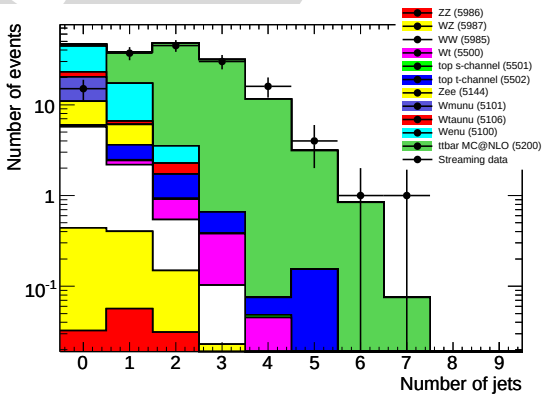
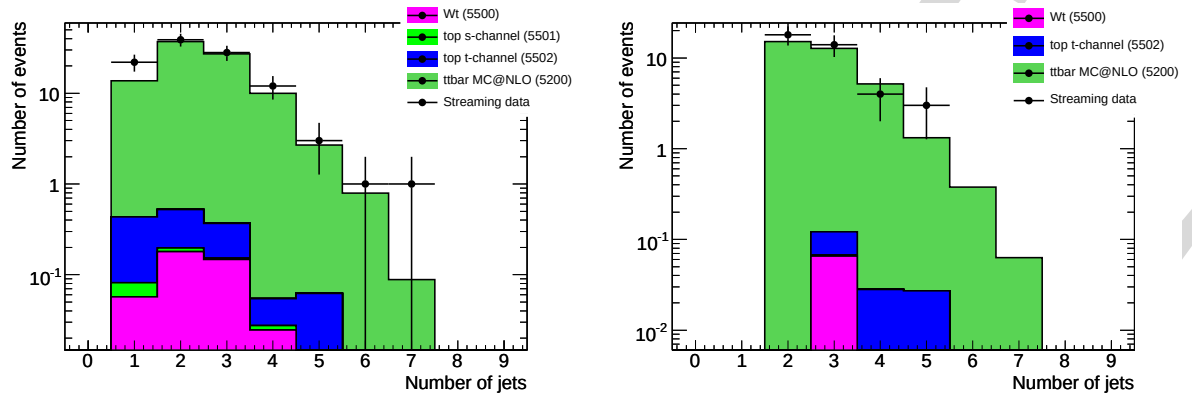


Figure 18: Normalized jet multiplicities for streaming data and its different components without b-tagging for fully leptonic event candidates.

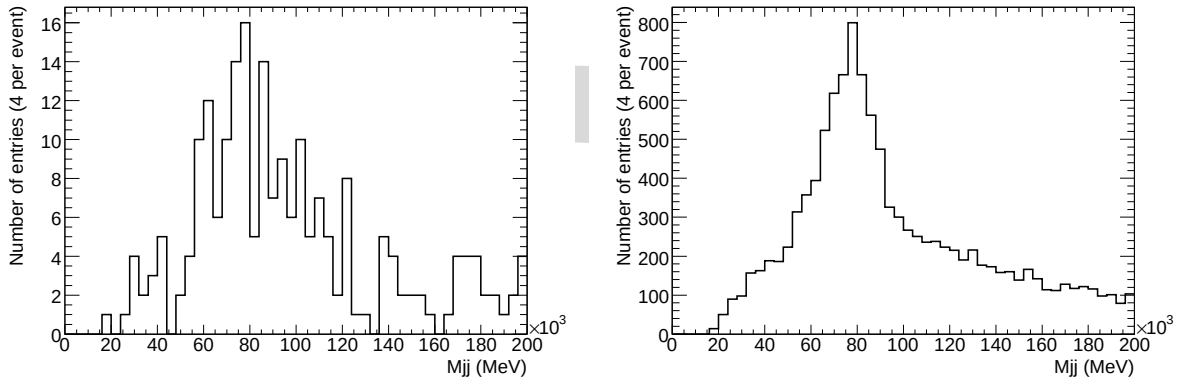


(a) (b)

Figure 19: Normalized jet multiplicities for streaming data and its different components when requiring at least one (a) or at least two *b*-jets (b) in the dilepton mode.

8.3 Reconstructed Top Mass

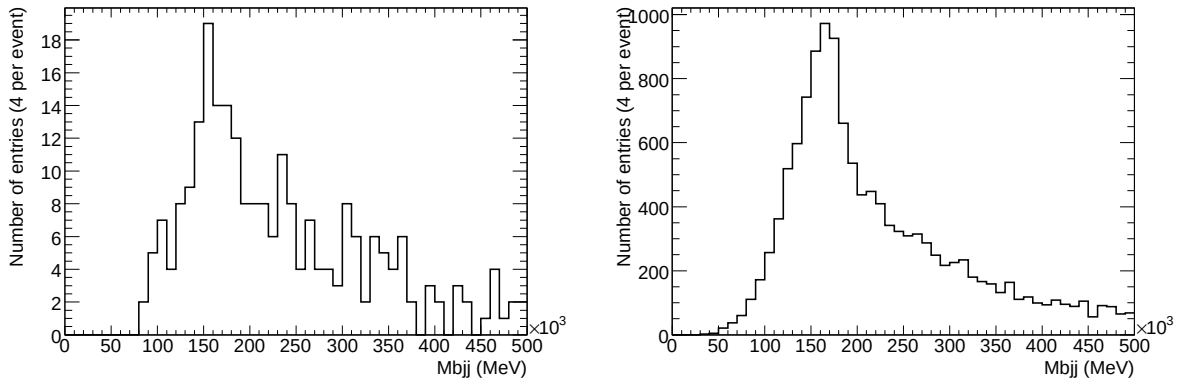
To validate the $t\bar{t}$ event selection for the semi-leptonic mode we consider the invariant mass of the hadronically decaying W and corresponding reconstructed top mass. We isolate events with four or five jets of which two are tagged as *b*-jets. For events with four jets, the hadronic W mass is calculated from the two untagged jets. For events with five jets there are three possible combinations of the W mass and for these we choose to have three entries per event. The distribution of the invariant mass is shown in Fig. 20 for streaming data and MC@NLO $t\bar{t}$ respectively.



(a) (b)

Figure 20: Invariant mass of the hadronic W for streaming data (a) and MC@NLO $t\bar{t}$ (b) for events with four or five jets of which two are *b*-tagged. There are 3 entries per event for 5-jet events.

The reconstructed top mass is determined by choosing the three-jet combination of the di-jets constituting the hadronic W together with the *b*-jet resulting in highest total transverse momentum. As for the W invariant mass we have one entry per event for the four-jet bin and three entries per event for the five-jet bin. The distribution of the reconstructed top mass is shown in Fig. 21 for streaming data and MC@NLO $t\bar{t}$.



(a) (b)

Figure 21: Reconstructed top mass for events with four or five jets of which two are *b*-tagged for streaming data (a) and MC@NLO $t\bar{t}$ (b). For each hadronic W di-jet combination, the three-jet combination (two untagged jets plus one *b*-jet) resulting in highest sum p_T is chosen.

9 Conclusion

Compare with [10].

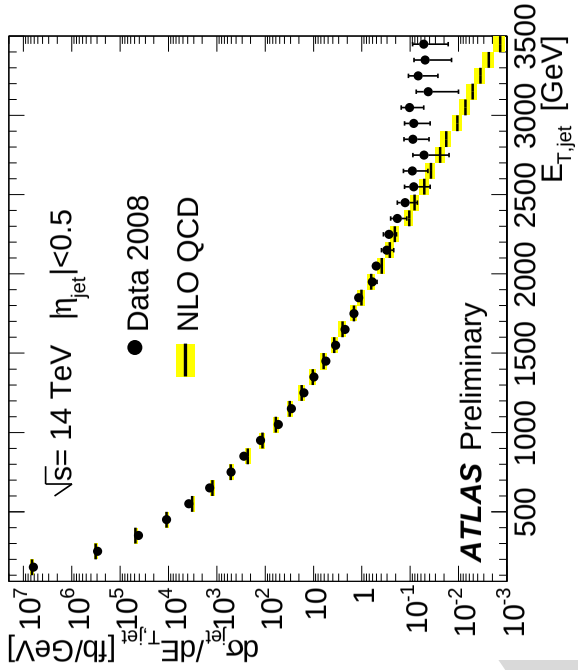
A Transfer function

Determining the systematic error on the backgrounds in Section **??** required that we generate high statistics samples with different Monte Carlo programs and parameter settings. We used a parameterization to represent the result of reconstructing jets and electrons in these samples, because ATLFast does not accurately model the smearing of these objects in the release 12 simulation.

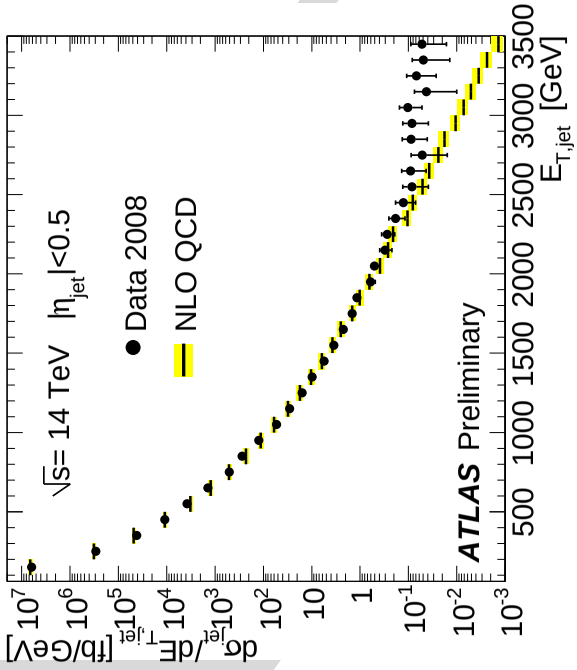
The purpose of the transfer function was to model the energy resolution and efficiency of reconstructed jets from the *Cone4TowerParticleJets* collection and electrons found by the EGamma algorithm, with respect to Cone4TruthParticleJets and **GenXX electrons with the** `isGenStable()`**}flag set. Since the primary application of the transfer functions is in inclusive** W **and** Z **samples similar to 6 - 7 in Appendix B, we used events from sample 7 (**PYTHIA **version XXXXX in release 12) to derive the transfer functions. This inclusive** $W \rightarrow e\nu$ **dataset is the same as our baseline** W **background sample.**

A.1 Truth jet transfer function

Truth jets were created from all stable outgoing particles in the mc collection, using the cone 0.4 jet clustering algorithm [?]. Profile histograms of the resolution $E_T(\textbf{truth}) - E_T(\textbf{reco})$ **as a function of jet** E_T **and** η **are shown in Figure 22. The efficiency to reconstruct these truth jets was measured by counting the fraction of truth jets that matched a reconstructed jet within a cone of 0.3 right?. The reconstructed jets considered for matching had a** *de facto* **energy threshold of 7 GeV. The efficiency curves are shown in Figure 23.**

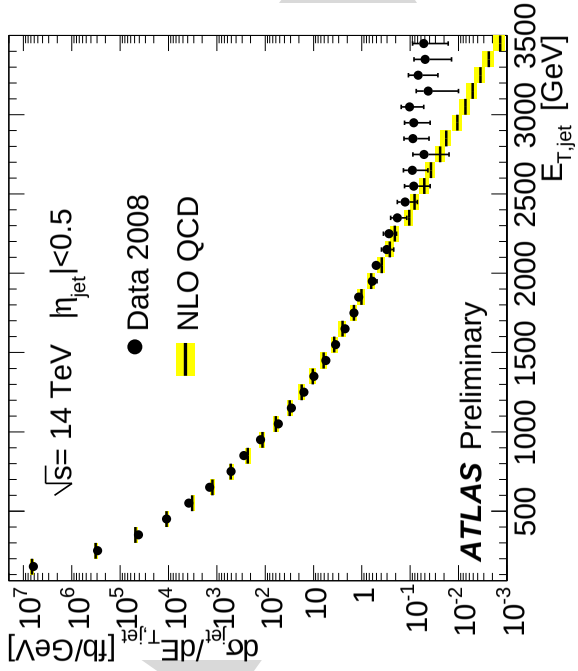


(a) Transfer function projection(?) with respect to jet E_T.

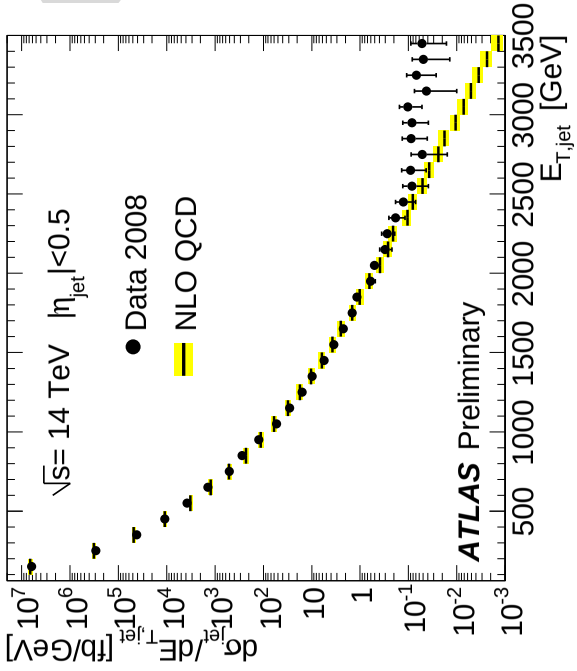


(b) Transfer function projection(?) with respect to jet η.

Figure 22: Transfer function for cone 0.4 jets derived from PYTHIA events (sample 7).



(a) Efficiency with respect to truth jet E_T.

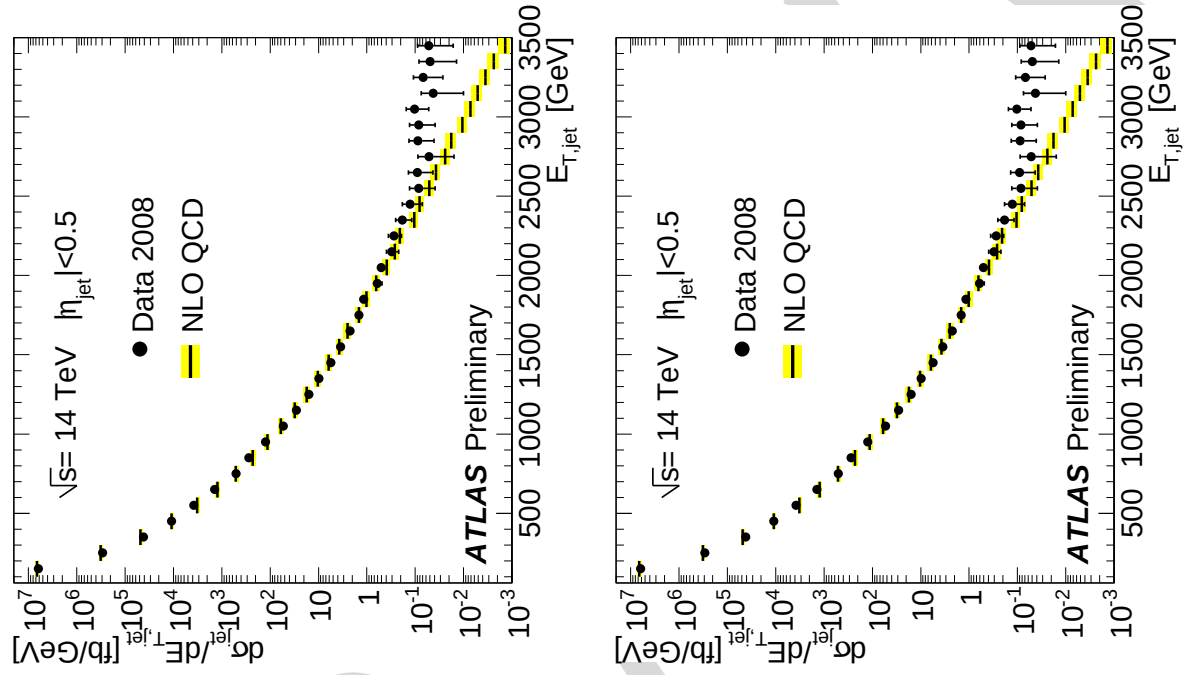


(b) Efficiency with respect to truth jet η.

Figure 23: Efficiency to reconstruct cone 0.4 tower jets derived from PYTHIA events (sample **??**).

A.2 Electron transfer function

The transverse energy of electrons from final electrons from hepMC was compared to reconstructed electrons passing all of the cuts in Section??. The reconstruction efficiency vs E_T and η is shown in Figure ?? and the resolution profile histogram is shown in Figure ??.



(a) Transfer function projection with respect to electron E_T. (b) Transfer function projection with respect to electron η.

Figure 24: Transfer function for electrons derived from PYTHIA events (sample **??**).

A.3 Transfer function systematics

We compare the transfer functions constructed above to transfer functions derived from other datasets, to determine if there is any generator dependence in the resolutions or efficiencies. The E_T dependence of $E_T(truth) - E_T(reco)$ distributions for jets from list of samples is shown in Figure A.3. Conclusion based on the plot.

B MC samples

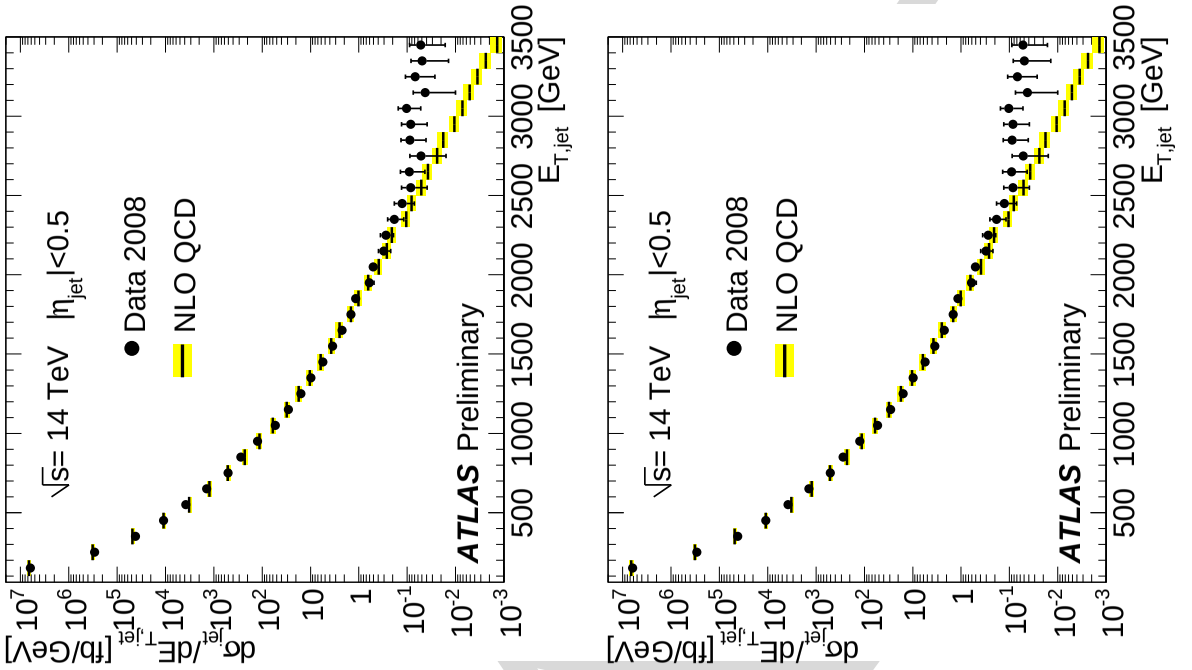
B.1 Samples for $t\bar{t}$ signal

MC@NLO : `trig1_misal1_mc12.005200.T1_McAtNlo_Jimmy.recon.AOD.v12000605` (2)

ACERMC : `trig1_misal1_mc12.005205.AcerMCttbar.merge.AOD.v12000604` **(bad!)** (3)

PYTHIA : `trig1_misal1_mc12.005568.ttbar_Pythia.merge.v12000605` (4)

(5)



(a) Efficiency with respect to truth jet E_T.

(b) Efficiency with respect to truth jet η.

Figure 25: Efficiency to reconstruct electrons derived from PYTHIA events (sample **??**).

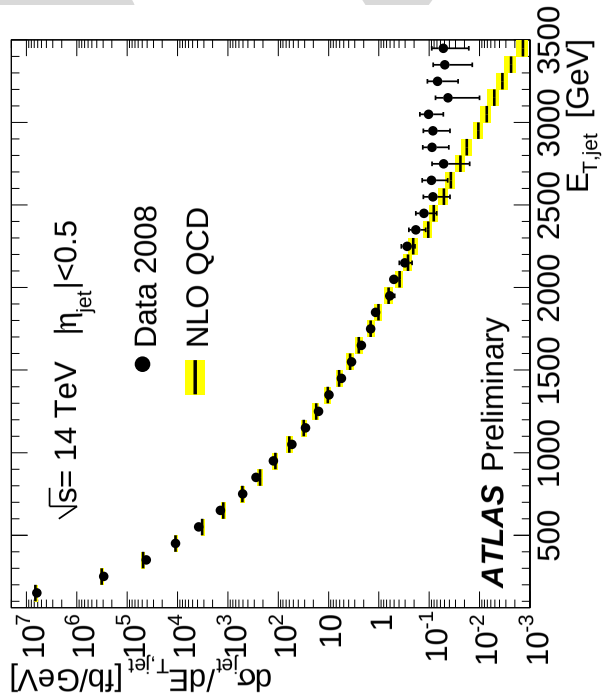


Figure 26: Efficiency to reconstruct electrons derived from PYTHIA events (sample **??**).

B.2 Samples for electroweak backgrounds

$W \to e\nu$**Jimmy** : `trig1_misal1_csc11.005100.JimmyWenu.recon.AOD.v12000601` (6)

$W \to e\nu$PYTHIA : `trig1_misal1_csc11.005104.PythiaWenu.recon.AOD.v12000601` (7)

$Z \to ee$: `trig1_misal1_mc12.005144.PythiaZee.recon.AOD.v12000604` (8)

$W \to \tau\nu$: `trig1_misal1_csc11.005106.PythiaWtaunu.recon.AOD.v12000605` (9)

(10)

B.3 Samples for single top and dibosons

$t-$**channel** : `trig1_misal1_mc12.005500.AcerMC_tchan.merge.AOD.v12000605` (11)

Wt : `trig1_misal1_mc12.005500.AcerMC_Wt.merge.AOD.v12000605` (12)

$s-$**channel** : `trig1_misal1_mc12.005500.AcerMC_schan.merge.AOD.v12000605` (13)

WW : (14)

WZ : (15)

ZZ : (16)

(17)

Describe generator filter for each sample.

References

[1] EtMiss, https://twiki.cern.ch/twiki/bin/view/Atlas/ETMiss.

[2] NewsForPhysicsUsers, https://twiki.cern.ch/twiki/bin/view/Atlas/NewsForPhysicsUsers.

[3] M. J. Flowerdew, "Electron fake rates and trigger efficiency,", http://indico.cern.ch/conferenceDisplay.py?confId=12054.

[4] S. Frixione and B. R. Webber, The MC@NLO Event Generator, 2002.

[5] G. Corcella and I. G. Knowles and G. Marchesini and S. Moretti and K. Odagiri and P. Richardson and M. H. Seymour and B. R. Webber, JHEP 0101 (2001) 010.

[6] A. Sen-Gupta and F. S. Merritt and J. Proudfoot and A. Farilla and M. Verducci, ATL-COM-PHYS-2007-031 (2007) 29.

[7] Alwall, J. and others, (2007).

[8] Kersevan, Borut Paul and Richter-Was, Elzbieta, (2004).

[9] , .

[10] B. Acharya et. al., TTbar X-section from StreamTest, http://indico.cern.ch/contributionDisplay.py?confId=22008.